IN THE UNITED STATES DISTRICT COURT
FOR THE DISTRICT OF COLUMBIA

ELOUISE PEPION COBELL, et al.,	)	
	)	
Plaintiffs,	)	
	)	
v.	)	Case No. 1:96cv01285(JR)
	)	
DIRK KEMPTHORNE,	)	
Secretary of the Interior, et al.,	)	
	)	
	)	
Defendants.	)	
	)	

**DEFENDANTS' FILING OF RESPONDING EXPERT REPORT
OF SUSAN HINKINS, PURSUANT TO RULE 26(a)(2)
OF THE FEDERAL RULES OF CIVIL PROCEDURE**

Defendants hereby file and attach hereto the Responding Expert Report of Susan Hinkins.

Dated: September, 17, 2007.

<div style="margin-left:50%">

Respectfully submitted,
PETER D. KEISLER
Assistant Attorney General
MICHAEL F. HERTZ
Deputy Assistant Attorney General
J. CHRISTOPHER KOHN
Director

 /s/ Robert E. Kirschman, Jr.
ROBERT E. KIRSCHMAN, Jr.
(D.C. Bar No. 406635)
Deputy Director
Commercial Litigation Branch
Civil Division
P.O. Box 875
Ben Franklin Station
Washington, D.C. 20044-0875
Phone (202) 616-0328
Fax (202) 514-9163

</div>

CERTIFICATE OF SERVICE

I hereby certify that, on September 17, 2007 the foregoing *Defendants' Filing of Responding Expert Report of Susan Hinkins, Pursuant to Rule 26(a)(2) of the Federal Rules of Civil Procedure* was served by Electronic Case Filing, and on the following who is not registered for Electronic Case Filing, by facsimile:

> Earl Old Person (*Pro se*)
> Blackfeet Tribe
> P.O. Box 850
> Browning, MT 59417
> Fax (406) 338-7530

>   /s/ Kevin P. Kingston
>   Kevin P. Kingston

Cobell v. Kempthorne

Response to Expert Report of Dwight J. Duncan

September 17, 2007

Susan Hinkins, Ph.D.
Hinkins-Susan@norc.org
Phone: 406-522-0164
Fax: 406-522-0165

Table of Contents

Appendices

Executive Summary

The statistical sampling and estimation procedures used in the Litigation Support Accounting (LSA) Project and the 2007 Plan are straightforward applications of statistical sampling to a target population. They are not critically flawed. In this section, I summarize the corrections and clarification made in this report to the inaccuracies in Mr. Duncan's Report.

(1) The LSA sample design was appropriate for its purpose

Mr. Duncan incorrectly states that "the statistical sampling design employed ***attribute sampling***, which is designed to answer yes or no questions such as *was money collected properly recorded*, and NOT to answer questions of accuracy such as *how much money was collected*."[1] He appears to believe that the intention of the LSA reconciliation process was only to flag a transaction as error/no error. In fact, the dollar difference was measured by the reconciliation and the sample was designed for this purpose. The sample was designed to measure both the amount of the dollar difference and the error rate. It is usually the case that a statistical sample is designed to address more than one type of information. Most surveys contain both "attribute" information (e.g., own a TV – yes/no) and "variable" information (e.g., the number of hours spent watching TV). The statistical issues are summarized below and described in more detail in Section 2 of this report.

- NORC designed the sample to provide information about the accuracy of the Individual Indian Monies (IIM) account **transactions** recorded in the electronic data systems.
- The LSA sample was appropriate for the purpose of estimating both the attribute (percentage of transactions in error) and the variable (the dollar difference).

(2) Inferences were made to the target population from which the sample was selected

Mr. Duncan mischaracterizes the LSA Project results as being applied to cover the entire Electronic Ledger Era. NORC's reports and the 2007 Plan clearly state that the target population for the LSA Project was restricted to transactions currently available in an electronic data base, and this does not include all transactions in the Electronic Ledger Era. Many of Mr. Duncan's criticisms are not valid because he does not recognize this distinction and he does not consider the other tests described in the 2007 Plan. Mr. Duncan challenges the use of sampling with the assertion that the tests and sample results described in the 2007 Plan cannot be extrapolated beyond the target definition used in the Plan. But if the decision is made to expand the population of interest, the statistical sampling can also be expanded accordingly. Section 3 of this Report addresses the following statements in more detail.

- The target population for the LSA Project consisted of approximately 28 million transactions recorded and identifiable in the IRMS and TFAS data base[2] at the time of sample selection. The results of the LSA Project are correctly applied to that

---

[1] Expert Report of Dwight Duncan, page 5 of 79.
[2] IRMS is the acronym for the Integrated Records Management System and TFAS refers to the Trust Fund Accounting System.

population. It is statistical evidence, not conjecture. These results are relevant and informative.

- These 28 million transactions **do not** constitute the entire population of Electronic Ledger Era transactions, as Mr. Duncan incorrectly claims.
- The 2007 Plan includes the posting test and additional supplemental sampling to specifically address other transactions from the Electronic Ledger Era that were not covered under the LSA Project.
- If the population of interest is expanded beyond that described in the 2007 Plan, the statistical sampling can also be expanded so that supportable inference can be made about the larger target population.

(3) NORC's LSA results use appropriate measures of "error"

Mr. Duncan's statement that all unsupported transactions in the LSA are treated the same as reconciled transactions is not correct. As documented in the NORC LSA report, omitted transactions and transactions without supporting documentation are **NOT** treated the same as supported transactions. Mr. Duncan's theory as to how the errors were counted is incorrect, as summarized below. Section 4 of this report describes how the error rate was calculated based on the data collected during the reconciliation process.

- The dollars in error were measured, recorded and used in the estimation process.
- Un-reconciled transactions were not "counted as reconciled and correct."
- NORC estimated the error rates by counting the **unreconciled** transactions as **errors.**
- There was absolutely no netting of errors between two transactions (an underpayment did not cancel an overpayment).

(4) The Qualitative Meta-Analysis Report is not central to the sample design

Mr. Duncan has misunderstood the purpose and use of the Qualitative Meta-Analysis Report. It is not central to NORC's sample design or estimation for the Paper Ledger Era, and to my knowledge, it is not central to the 2007 Plan. This issue is addressed in my responses in Appendix A and more fully in Dr. Scheuren's rebuttal report.

1. Introduction

This is the first of two rebuttal reports that the National Opinion Research Center (NORC) has prepared to respond to comments that Plaintiffs' expert Mr. Duncan[3] has made. In this report, I correct errors or misstatements made by Mr. Duncan regarding statistical issues in his Export Report of August 23, 2007. In the second NORC rebuttal report, Dr. Scheuren responds to Mr. Duncan's claims that the Qualitative Meta-Analysis is misleading.

I am a co-leader with Dr. Fritz Scheuren for NORC's engagement with the Department of the Interior's Office of Historical Trust Accounting. I have also been the project manager since September 2002. I attended the meetings for the LSA Project at which the scope and objectives were discussed and the procedures for the LSA Project were planned. Dr. Scheuren and I directed NORC's work on the sample design, sample selection, and estimation. We have also participated in planning sessions for the remaining samples for the Electronic Ledger Era and the future work to test the Paper Ledger Era.

Mr. Duncan's report contains substantive errors or misstatements regarding the sample design, estimation and application of the LSA Project results. My report is organized around the following three general statistical topics addressed by Mr. Duncan.

- Design objectives and sampling to estimate attributes and amounts (Section 2)
- Applying sample results to the target population from which the sample is drawn (Section 3)
- Definition of error and calculating estimates (Section 4)

The fourth statistical topic addressed by Mr. Duncan is the use of qualitative meta-analysis. While I address this issue in Appendix A, Dr. Scheuren's Rebuttal Report provides NORC's detailed response on meta-analysis.

Appendix A provides my page-by-page response to specific statements made in Mr. Duncan's Report.

2. Sample Objectives, Sample Design and "Attribute sampling"

This section provides additional information for the following general points which I feel need to be clarified.

- NORC's understanding was that the objective of the sampling was to provide information about the accuracy of the **transactions** recorded in the electronic data bases, available at the time of sample selection.
- The reconciliation process collected the dollar differences for the sampled transactions.
- The sample used was appropriate for the purpose described in the first bullet above, and it was designed to estimate both the attribute and the "variable", i.e., the dollar difference.

---

[3] Referenced in the August 23, 2007 report to the court entitled *Expert Report of Dwight J. Duncan, CFA*.

The first step for the statistician is to understand the objectives of the sample and to determine what information should be collected and recorded about the sampled units. NORC understood that the primary purpose of the sample for the Litigation Support Accounting Project (LSA) was to "shed light on the accuracy of the land-based Individual Indian Monies (IIM) account transactions contained in the two IIM Trust electronic systems – Integrated Records Management System (IRMS) and Trust Fund Accounting System (TFAS)."[4] It was also my understanding that this information could be used to provide the accountholders with a measure of assurance regarding the accuracy of the transactions in the LSA target population. Accuracy was to be estimated both in terms of the proportion of transactions with error and in terms of the dollars in error.

Because the population of transactions was very large, over 28 million transactions, sampling was used to provide estimates of accuracy, in terms of the <u>variable</u> (amount of dollars in error) and the <u>attribute</u> (whether or not an error occurred). The reconciliation process would measure the dollars in error for each reconciled transaction.

Once the objectives and the reconciliation process were determined, the next general step was to determine the appropriate target population and the "list" or "frame" that includes every population element. This is a crucial step in the sample design and it is discussed in detail in Section 3.

The next step was to design the sample. Available population data were analyzed and subject matter experts were consulted to understand what factors may be important to consider in the sample design. The statistician attempts to identify factors that can be used to reduce the variability in the estimate, by identifying similar transactions, and to ensure that important characteristics are covered within the sample.

Coverage of important characteristics was NORC's primary design concern in the LSA Project. Stratification is a commonly used tool in designing samples as it ensures that the sample covers important characteristics of the population. The LSA design was stratified by time periods, to ensure coverage over different processing systems, by the size of the transaction, the location (BIA Region), and by the type of transaction.

A key decision in the sample design is the determination of the sample size. Many factors are considered in making this determination, including the need to cover the important variables in the population using stratification, as discussed above. Another factor to consider is the likely outcome of the sample, and in particular the expected variability that might be observed

Before the reconciliation is completed, it is difficult to know or even guess at the size and in particular the variability of the dollars in error. Therefore, it is standard statistical practice to make calculations relating sample size to precision and assurance, based on the attribute. Because of the nature of the probability calculations for attributes, when one makes an assumption about the proportion (e.g. if we assume a 2% error rate will be observed), this

---

[4] Reconciliation of the High Dollar and National Sample from Land Based IIM Accounts (All Regions), NORC, September 30, 2005, page 5

assumption and the population size is all that is needed to calculate the sample size needed to obtain specific assurance and precision.

Calculations such as those shown in the following table are often made to aid in the determination of the sample size.

Example of sample size calculations 99% assurance
(Population of 50,000 elements)

Observed Error Rate	Precision of the Estimate	
	5%	**1%**
0%	90	460
1%	140	1,000
5%	200	2,580

There are three key factors in this calculation: the expected observed error rate (first column), the desired assurance (99% in this table) and the desired level of precision (either .01 or .05 in this table.) The precision of the estimate measures the distance from the observed value to the upper bound. Consider the row in the table showing the calculations assuming 1% of the sample is in error. The point estimate would be 0.01 or 1%. With a sample of size 1000, one could make the 99% assurance statement that, having observed a 1% error rate, the true error rate is no more than 2% (the point estimate of .01 plus the "precision" of .01). If a sample of 140 had been used, the strongest 99% assurance statement would have been that the true error rate was no greater than 6% (point estimate of .01 plus the "precision" of .05). Larger sample sizes provide better precision.

The table also demonstrates the effect of the observed error rate. To obtain the same precision and assurance, larger sample sizes are needed for larger observed error rates (up to a 50% error rate).

Such calculations were done during the design of the LSA sample, but the sample size was not determined using only this attribute calculation. The primary concern in NORC's design was to ensure the sample covered known factors that might affect the size and likelihood of errors. A sufficiently large sample was required in order to allocate sample units across all of the desired strata, as described earlier.

The estimates were calculated by weighting the sample units by the number of transactions each sample unit represents. For example, sample units in a stratum where 5 units were sampled from a population of 100 units would have a weight of 20 – each sample represents 20 units. The sample weights ensure that the estimates are not biased by any disproportionate selection probabilities.

Because it was designed to produce data-supported conclusions regarding both the likelihood and the size of errors, the LSA sample design was not an "attribute sample." The dollar difference was measured by the reconciliation process. The design stratified the sample so that if either the likelihood or the size of a difference was correlated with the known, stratifying variables, this effect would be identified. The sample size was large enough to provide coverage across all such strata.

I will make one final comment regarding sample size. As briefly described above, the statistician must use available information to make an educated guess as to the necessary sample size to obtain the desired precision and assurance. If the sample is much too large for the stated objectives, resources are used that could have been better applied elsewhere. If the sample is much too small, the sample may not provide the assurance and precision required for the stated objective. But if the sample size is too small, this can be alleviated by increasing the sample size in a statistically appropriate manner.

Finally, a safeguard used in the LSA Project is that only the **upper** bound on the error was estimated (e.g. "a 99% upper bound on the dollar exposure for debit transactions is estimated to be under $4 million"[5].) This means that if the estimate has a greater sampling variability than desired, e.g., a small sample size, the result would be to increase the upper bound, estimating a larger error.

3. Target Population and Sampling Frame Concerns

Mr. Duncan repeatedly asserts that the results of the LSA Project were applied to the entire target population of Electronic Ledger Era Results. This is incorrect and this erroneous assumption results in additional incorrect or misleading statements by Mr. Duncan. This section discusses the following points.

- The results of the LSA Project are correctly applied to a population of approximately 28 million transactions.
- The 28 million transactions in the target population tested by the LSA Project do not and were not intended to constitute the entire population of Electronic Ledger Era transactions.
- The 2007 Plan describes additional tests for other portions of the population of Electronic Ledger Era transactions.
- In particular, the LSA Project does not address transactions that were never recorded in the government electronic records. By its very design, the LSA Project could not address this source of error and in the LSA estimates such "missing" transactions are not counted as "accurate".

3.1 Applying Sample Results to the Correct Population. Statisticians will generally agree that the results of a sample should only be applied to the target population from which it was selected. To apply the results to a broader or a different population requires assumptions or additional knowledge.

Mr. Duncan mischaracterizes the LSA Project results as being applied to the entire Electronic Ledger Era. NORC's reports and the 2007 Plan clearly recognize that the target population for the LSA Project does not include all transactions in the Electronic Ledger Era. Furthermore, the 2007 Plan describes tests that will be done to address portions of the population not included under the LSA Project.

---

[5] LSA report page 14.

It is important to recognize that the sampling tests that, to-date, NORC has designed samples can test only those transactions for which there is some evidence within the DOI records, i.e., transactions that were recorded in the IIM accounting system (electronic or paper ledgers) or transactions for which there is evidence in other government records (e.g. computer printouts or entries in a lease log). The target populations available to us are defined by information available in the DOI systems, but this information is not limited to electronic data bases or to ledgers.

One of the first tasks for the statistician is to determine the client's population of interest (i.e., the target population) and whether there is a complete and definitive list of this population from which to sample. For a statistically sound result, each member of the population must have a chance to be selected. The development and testing of the frame from which the sample is drawn is a crucial step because the usefulness of the sample depends on it.

In a project as large and complex as that undertaken by the OHTA, it is often necessary to divide the work into manageable tasks. Different procedures may be required or may be more efficient for different portions of the population. A pilot test of new procedures is often valuable and doing work in stages allows one to apply the lessons learned from one sampling effort to improve the efficiency in another sampling effort. It can be statistically valid to report on portions of the testing that have been completed before all sections have been completed. Of course it is very important to restrict the conclusions to the portion of the population that has been tested.

For example, the sample selection and reconciliation for the Eastern Region was completed in advance of the others. It was appropriate and statistically valid to make estimates and assurance statements about the population of transactions in this one Region. In the Eastern Region Report, NORC did not apply inferences from the Eastern Region to all Regions. But the information for this one Region may have been interesting or useful to the client. When the samples over all Regions were completed, these data, including the Eastern Region data, were used to make inference about the LSA target population.

Similarly, NORC's LSA Report very specifically defines the target population to which the inference applies. While the LSA target population is not the entire target population of all transactions in the Electronic Ledger Era, it covers a large number of transactions (28 million). The 2007 Plan lists the tests to be used to address the "gaps" and other portions of the population not covered by the LSA Project. These are summarized in the following section.

Finally, Mr. Duncan states that the statistical sampling in the 2007 Plan does not provide a basis for extrapolation beyond its target population.[6] I agree. As described in this section, the statistical sampling is designed to address the target populations and the objectives stated in the 2007 Plan. However, if the decision is made to enlarge the target population beyond the definition used in the Plan, then the statistical sampling can also be expanded so that supportable inference can be made.

---

[6] Bottom of Page 5 of 79 of Mr. Duncan's Expert Report.

3.2 Partitioning the Target Population of Transactions in the Electronic Ledger Era. This section briefly outlines the partitioning of the population of IIM transactions and the test to be applied in each partition. I will confine my comments to addressing only the Electronic Ledger Era transactions, in land-based IIM accounts that were open, either as of December 31, 2000, or on or after October 25, 1994, but closed prior to December 31, 2000. Also, as described in the previous section, the **sampling** tests are applied to those transactions for which there is some evidence within the DOI systems and other records.

With these caveats, the following table summarizes my understanding of the partition of the transactions and the proposed test for each partition.

Transactions in the Electronic Ledger Era

Defined Partition of Population	Test
1. Receipt transactions never recorded in the government ledgers	Posting Test
Transactions recorded at some time in the government IRMS or TFAS system	
2. Interest transactions	Interest Recalculation
3. Ledger balance transfers	One of the Paper Ledger Era Tests
4. Non-interest, non-ledger-balance transactions **available** in the electronic record and **identifiable** as being in the target population, at the time of sample selection[7]	LSA Project
5. Transactions **available** in the electronic record at the time of sample selection but **not identified** at that time as being in the target population	"Follow-up sample" after this portion of the population is identified by Interest Recalculation and/or Data Completeness Validation (DCV) test[8]
6. Transactions **not available** at the time of sample selection because, while they were recorded electronically at the initial conversion to IRMS, they were subsequently over-written or otherwise deleted	One of the Paper Ledger Era Tests
7. Transactions, subsequent to the IRMS conversion, which were **not available** in the electronic data at the time of sample selection	"Follow-up sample" after this portion of the population is identified by DCV

First, transactions that were **never** recorded in either the paper ledgers or the electronic ledgers cannot be tested by sampling from these ledgers. Therefore, to provide information regarding the completeness of the transactional history, the 2007 Plan includes the Posting

---

[7] For the LSA Project, the 10-region sample was selected in December of 2003. The Eastern and Alaska region samples were selected earlier.
[8] DCV includes among many tests, mathematical recomputation of posted transactions to determine the arithmetic total of transactions posted between time periods (e.g., month ends). If the transactions do not recompute, a transaction may have been omitted from the database. A search of records is commenced to locate the missing transaction.

10

test which traces receipts (selected independently of account statements) to accounts. This is the first partition in the table above.

The remainder of the table addresses transactions that were recorded at some point in the electronic ledgers. Here, one partition is defined to include interest transactions which will be tested by the interest recalculation, and the second contains the ledger balance transfers, which will be tested with the Paper Ledger Era transactions.

For the remaining transactions (non-interest, non-balance transfer transactions that were recorded in the electronic systems at some time), the LSA Project still does not provide complete coverage. The LSA Project tested only those transactions recorded in the IRMS and TFAS systems and for which the data **were available and identifiable electronically at the time of selection** (the fourth partition).

Partitions 5 - 7 describe the remaining categories of transactions which are not covered by the LSA Project. Ideally the order of the process would have been to first complete the Data Completeness Validation (DCV) and the interest recalculation. At the time that the sample was selected, it was known that the DCV and interest recalculation were likely to identify transactions as described in partitions 5 – 7 and these transactions would need to be tested separately, after a list covering this portion of the population was available.

Partition 5 includes transactions available in the data but not identifiable as being in the target population, at the time of selection. For example, the interest recalculation will identify any non-interest transactions that were incorrectly excluded from the sample frame because they were coded as "interest." Because the subsequent processes will identify transactions that should have been included, these transactions can be tested with follow-up sampling at that time.

The remaining two classes contain transactions which were not available in the data file at the time of sample selection, i.e., 'gaps.' In one case (6), the transactions are on the "boundary" between the "Electronic" and the "Paper." It is my understanding that testing these transactions will require research in the "Paper Ledger Era," and so these transactions will be tested with the Paper Ledger Era tests. In Section 4.3.1 of Mr. Duncan's Report, he refers to these transactions as a "vast quantity of missing data in the Electronic Ledger Era". As stated here, while these transactions are not tested under the LSA Project, because they were missing from the currently available electronic data, it is my understanding that the information is available from paper records and these transactions will not be "missed" but will be tested in the work still remaining.

The last case covers the data gaps after the inception of IRMS. It is my understanding that in most cases, the DCV work will obtain the transactional information from the paper documents. In these cases, the transactions will be made available in an electronic data base and can be tested via the follow-up sample tests, described in the 2007 Plan. If there are gaps where the transactional information cannot be identified, my recommendation would be that such gaps would be documented, reported, and estimation developed through additional tests or conservative models, as discussed in Section 4.

The LSA Project results do not cover the data "gaps" described above, and NORC did not claim otherwise. The issue of coverage by the sample frame was clearly addressed in the Limitations section of NORC's report:

> *In addition, the sample was drawn from electronic data files available in December 2003 before data validation was completed and prior to interest recalculations. While it would have been ideal to have waited to complete these two procedures prior to sampling, this was not practical. These procedures are underway, and they have each identified transactions that were not included in the electronic data file. These are transactions that are in the accountholders electronic account statement but they have not been tested in the process described here nor tested in other ways. This under-coverage issue can be addressed by reconciling a random sample of transactions drawn from these identified transactions.*

Transactions not included in the sampling frame (e.g. "gaps") were not counted as reconciled and complete. They were not 'counted' or estimated at all in the NORC results. The transactions not covered by the Project were specifically identified and the 2007 Plan describes the tests which will be done for these transactions. Estimates for these partitions of the population cannot be made until the additional tests are completed.

The LSA sample results are correctly applied to the 28 million transactions in partition #4. The error in Mr. Duncan's report is to assert that NORC or the 2007 Plan claimed that these 28 million transactions include all transactions in the population of interest.

One of the statistician's responsibilities is to ensure that the sample results are correctly applied to the population from which it was drawn. NORC has always informed DOI as to exactly what the sample does and does not cover, and we will continue to do so.

3.3 Out-of-Scope Transactions. Because Mr. Duncan, in his Report, articulated concerns about this process, the remaining portion of this section addresses out-of-scope transactions.

Rarely does one find a perfect sampling "frame" or "list" for the target population. The first concern is that every element in the population can be selected from the sampling list. This concern was addressed in Section 3.2.

The second concern is whether the sampling frame includes elements that are not in the target population. Because the primary concern is to cover the target population, it is quite common to use a sampling frame that covers more than the target population in order to ensure that no population unit is excluded. While extraneous elements affect the efficiency of the sample, they do not bias the results.

The NORC LSA Report describes the types of transactions that were intended to be excluded from the target population – e.g. transactions in administrative accounts, interest transactions, and ledger balance transactions. However, it was not always possible to identify such elements from the data available at the time of sample selection. In such cases, it is better to include elements that may be out-of-scope rather than incorrectly exclude elements that are in the target population. This is a common characteristic of sample frames and therefore the

definition of "out-of-scope" or "out-of-population" transactions is well-known to sampling statisticians.[9]

It is the statistician's responsibility to account for each selected unit and to ensure that no sampled unit is "ignored." Every out-of-scope transaction was determined to be either:

1. excluded from the population of interest completely (e.g. transactions in administrative accounts) or
2. contained in a different partition and tested in another portion of the Plan (e.g. interest transactions).

These transactions were not ignored but were handled according to correct statistical procedures. In some cases, all out-of-scope elements can be identified and removed from the population. Usually, however, all such elements cannot be identified in the population and the in-scope population size must be estimated. The effect of the out-of-scope transactions was shown in Table 1 of NORC's LSA Report, and one can see that the out-of-scope transactions were a minimal part of the sample.

The out-of-scope transactions do not enter into the estimation for the target population, but they are retained in the sample data in the Account Reconciliation Tool (ART)[10]. Appendix B in NORC's LSA Report provides specific information on the transactions that were found to be out-of-scope and shows explicitly how the in-scope target population was estimated based on these transactions.

4. Definition of Error and Calculation of Estimates

There are serious misstatements made by Mr. Duncan about how the reconciliation in the LSA Project measured errors, how errors were counted, and how the estimation was done. This section emphasizes the following points.

- The dollars in error were measured, recorded in the data base, and used in the estimation process.
- There was absolutely no netting of errors between two transactions (an underpayment did not cancel an overpayment).
- Un-reconciled transactions were not counted as "reconciled and correct." As clearly stated in NORC's LSA Report, NORC estimated the error rates by counting the **unreconciled** transactions as **errors.**

This section describes the procedures used for defining errors and calculating estimates. The following information, except the tabulation by accounting codes, was included in NORC's LSA Report.

---

[9] Briefly, an out-of-scope transaction is a transaction that was included in the sampling list (frame) but which is not included in the definition of the target population. For example, there were transactions selected from the frame that when examined by the accountants, were determined to be interest transactions or ledger balance transfers. Upon examination of one account, it was found to be an administrative account, not an IIM account, and therefore not contained within the target population.

[10] The Account Reconciliation Tool (ART) is a data management system that is used to record the reconciliation results.

4.1 Data returned by the reconciliation process. Four accounting firms were contracted to reconcile the selected transactions. The reconciliation was performed according to the Accounting Standards Manual (ASM). A fifth accounting firm performed quality assurance tests to ensure that the ASM procedures were followed.

For all Regions except the Eastern Region, the accountants entered the reconciliation information directly into the centralized data base on the ART. From this information, two data items were used by NORC for estimation. First, the accountants indicated whether or not sufficient supporting documents had been located in order to reconcile all aspects of the transaction, as defined in the ASM. This information was provided in the data field referred to as the accounting code. This is critical information since an unreconciled transaction cannot provide information on the dollars in error for that transaction.

The accountants also entered all dollar differences between the recorded transaction amount and the supporting documentation, for each reconciled transaction. The accountants did not enter an attribute indicator of "error" or "no error." The data base contains both the reconciliation status and the dollar differences, from which the attribute of whether or not there was an "error" was determined.

4.2 Reconciliation Status. The procedure for assigning the different values for the accounting code is defined in the ASM. The following reflects my understanding of the meaning of the accounting code values, as they pertain to using the information for estimation.

- Accounting codes 1 and 2 indicate the transaction has been reconciled to the standards of the Accounting Standards Manual (ASM).
- Accounting code 3 indicates that supporting documents were located that allowed the accountants to test some parts of the process, but because one or more of the pertinent documents were not located, the transaction cannot be considered fully reconciled.
- Accounting code 4 is assigned when little to no determination about the accuracy of the transaction could be made.

It is my understanding that transactions were submitted with accounting codes 3 or 4 when the accounting firm believed that every reasonable search had been made for the supporting documents.

At the time that the estimation was performed, there were also transactions selected but not submitted by the accounting firms. These are differentiated from the transactions with accounting codes 3 or 4 in that the accounting firm had not exhausted all reasonable search avenues for locating documents. Instead, "time" was called before the research could be finished. As described in Section 4.4, NORC's estimates counted these transactions as "errors."

The reconciliation of the Eastern Region sample was performed before the data base was developed in the ART. It is my understanding that when the Eastern Region transactions were reconciled, the option of indicating a "partially reconciled transaction" was not available. That is, the Eastern Region transactions were returned as

- reconciled directly (Accounting Code = 1),
- reconciled using alternative procedures (Accounting Code = 2), or
- not reconciled (Accounting Code = 4)

It was NORC's understanding, again, that both of the first two categories should be considered fully reconciled.

4.3 Estimation with incomplete data. A transaction which has not been reconciled provides no information as to whether or not there would be a difference if the transaction were reconciled. Only a person inexperienced in statistical estimation would make the mistake of treating "no information" as "no difference." This would not be good statistical practice. **NORC did not make this mistake**.

When sample data are incomplete, the statistician must use model assumptions to address the incomplete information. One model that can be appropriate in certain situations is to assume that the reconciled transactions are simply a random subsample of the original sample. Under this model, the assumption is that the un-reconciled transactions would have the same error rate as the reconciled transactions. This estimation technique may result in under-estimating the error rate if in fact the un-reconciled transactions are more likely to be in error.

As stated in NORC's report[11], estimates for the LSA Project were made using a conservative model designed to be more likely to over-estimate the error rate, than to under-state the errors. NORC calculated the estimates for the LSA target population by **assuming that the unreconciled transactions are in error**. This is a conservative estimate in the sense that this method is likely to over-estimate the true error rate. This estimation is described in more detail in the following section.

4.4 Estimation for the LSA Sample

The following tables show the number of transactions in the LSA in-scope sample (transactions under $100,000) by their reconciliation status at the time of estimation. The debit and credit transactions are tabulated separately. In each case, the first table shows the data by accounting code and the second table summarizes the same information using NORC's classification by reconciliation status.

Sampled Debit Transactions by Accounting Code

Reconciliation Status	Number of Transactions	
	11 Regions	Eastern Region
Submitted with Accounting Code = 1	2,176	109
Submitted with Accounting Code = 2	64	14
Submitted with Accounting Code = 3	0	-
Submitted with Accounting Code = 4	0	5
Not submitted	4	0
Total	2,244	128

---

[11] NORC's LSA Report, pages 14 and 17

Summary of Sampled Debit Transactions

Reconciliation Status	Number of Transactions
Fully Reconciled	2,363
Partially Reconciled	0
Not Reconciled	9
Total	2,372

Therefore, supporting documentation was found to reconcile 2,363 out of 2,372 selected debit transactions, for a completion rate of 99.6%. These are excellent results.

In the sampled debit transactions, no differences of $1 or more were found. If NORC had in fact counted the un-reconciled transactions as 'reconciled and accurate', the point estimate for debit transactions would have been zero. NORC's point estimate for the error rate is not zero but 0.4%.

NORC counted the 9 un-reconciled transactions as errors. This was reported in NORC's LSA Report, page 14, when referring to the 0.4% point estimate for debits: "The above figures are based on a very conservative estimate of the effect of the unreconciled transactions by calculating the estimates assuming that if reconciled, these transactions would have differences."

Sampled Credit Transactions by Reconciliation Code

Reconciliation Status	Number of Transactions	
	11 Regions	Eastern Region
Submitted with Accounting Code = 1	1,787	132
Submitted with Accounting Code = 2	167	28
Submitted with Accounting Code = 3	3	-
Submitted with Accounting Code = 4	1	1
Not submitted	9	0
Total	1,967	161

Summary of Sampled Credit Transactions

Reconciliation Status	Number of Transactions
Fully Reconciled	2,114
Partially Reconciled	3
Not Reconciled	11
Total	2,128

Therefore, supporting documentation was found to reconcile 2,114 out of 2,128 selected credit transactions, for a completion rate of 99.3%. Again, these are excellent results.

In the reconciled transactions, 36 differences of $1 or more were reported.

The estimation for credit transactions was complicated by the three transactions with partial information. Three estimation models were considered.

Using the model that the un-reconciled transactions are no different than the reconciled transactions, the estimate would be calculated based on only the 2,114 reconciled transactions and the weighted estimate of the difference rate would be 1.0%. This is not the estimate NORC reported.

NORC used the conservative assumption that the 11 transactions which were either not submitted or submitted with accounting code 4 would all have errors, if reconciled. This resulted in the estimated (weighted) error rate of 1.3%, as reported in NORC's LSA Report. I believe that this estimate is likely to over-estimate the true error rate.

An even more conservative estimate would be to assume that the three transactions with partial information were also in error. These partially reconciled transactions, however, were different from the transactions with no information – some information was available on the part of the process that could be tested with available documents.[12] Therefore, for the three transactions with partial information, NORC decided to use the information from the reconciliation, rather than to assume that they would have an error, and NORC counted the eleven transactions with no information as errors.

<u>4.5 Differences were not netted across transactions</u>. Section 4.3.3.1 of Mr. Duncan's report states "if two transactions in different accounts are misstated, one an over-payment of $100 and the other an under-payment of $100, the mean under-payment, according to DOI will be zero even though both of the transactions are misstated and both of the account balances are wrong."[13]

Mr. Duncan's implication in the fifth paragraph of his section 4.3.3.1 is that the estimates will hide errors by netting the overpayments with the underpayments. I want to emphasize here that no "netting" was used in the LSA estimation. The estimates provided were estimates of total error, that is total overpayments PLUS total underpayments. Estimates of total underpayments were also provided.

To illustrate how errors would be counted and the dollars accumulated, suppose that 100 transactions were reconciled with 50% of the transactions having an underpayment of $5 and 50% of the transactions having an overpayment of $5. NORC would report the results as having a 100% transaction error rate, not zero, as Mr. Duncan suggests. The total dollars in error would be reported as $500 with an underpayment of $250 and an overpayment of $250. From this information, the reader can calculate, if desired, that the net is zero.

---

[12] Difference amounts can be reported for a transaction with accounting code 3 and these transactions were subjected to the same ASM rules and the QA testing.
[13] Page 23 of 79.

5. Concluding Comments

Mr. Duncan's report consistently contains his claim that

> *DOI's consultants in this matter have employed certain statistical sampling procedures that are critically flawed and would likely confuse and mislead the lay person. Furthermore, the results of the statistical sampling procedures are inappropriately applied, and hence will not support the DOI's stated objectives.*

Mr. Duncan does not supply evidence that the sampling procedures are flawed and misleading. As one of the lead statisticians on NORC's team that provides statistical support services to OHTA, I would assert that NORC's sample design, data analysis and estimation techniques are not flawed but follow established statistical practices. NORC's practice is to use valid statistical practices that will "let the data speak."

Respectfully submitted,

_____

Susan Hinkins, Ph.D.

Appendix A. Section-by-Section Responses to Duncan Report

In what follows, I respond to Mr. Duncan's many claims that the statistical techniques that NORC has employed are flawed. The format is that direct quotes from Mr. Duncan's report will be *italicized*. My response to his accusations and misstatements will not be italicized. If Mr. Duncan added emphasis to his writing, I left the emphasis alone. I have not added any emphasis of my own to Mr. Duncan's writing. Mr. Duncan's footnotes are not included here, but the footnote marks in the main text have been maintained. If Mr. Duncan quoted other sources in his report, I have changed the font to make it clear that the quotes are not Mr. Duncan's.

Mr. Duncan's Executive Summary

Mr. Duncan - Page 5 of 79

DOI's consultants in this matter have employed certain statistical sampling procedures that are critically flawed and would likely confuse and mislead the lay person. Furthermore, the results of the statistical sampling procedures are inappropriately applied, and hence will not support the DOI's stated objectives.[7]

Response:

Mr. Duncan is incorrect in his statements regarding the statistical aspects of the LSA Project and the Plan. The statistical sampling and estimation procedures used are straightforward applications of standard statistical sampling to a target population. They are not critically flawed. The methodology is complex because the data we are dealing with are complex. While we do not expect a lay person to understand the complexities of the details of our analysis, NORC believes that it has explained the general concepts of the statistical analysis in a way that is understandable to a lay person. The overall standard for the analysis that NORC applied was to use established statistical techniques, i.e., ones that are well documented in the statistical literature and that will let the data speak.

Mr. Duncan - Page 5 of 79

The fundamental failings of both the historical transaction records as well as potential supporting documents due to missing and/or destroyed data has been substantiated by DOI's own accounting and statistical consultants (see Section 4.3.1.). Therefore, drawing a sample from the target population is not possible (see Section 4.3.2.).

The sampled population is different from the target population due to missing or omitted transactions; therefore any statistical inference from the 2007 Plan's sample to the target population is conjecture, not statistical inference (see Section 4.3.2.).

Response:

The results from the LSA Project were applied only to the target population for the LSA Project and are therefore appropriate statistical inference. The target population for the LSA Project was transactions recorded and identifiable in the IRMS and TFAS data base at the

time of selection. A valid sample from this target population was drawn and supporting documents were found for over 99% of the transactions. The LSA results can be correctly applied to over 28 million recorded transactions from which the sample was drawn. Section 3 in my report addresses this issue.

The target populations from which samples can be drawn by NORC are currently restricted to transactions for which there is some evidence within DOI records. But such evidence is not limited to electronic and paper ledgers and there are other sources of information such as computer printouts, lease logs, and contracts.

Mr. Duncan - Page 5 of 79

The statistical sampling procedures were designed for estimating litigation exposure and NOT for substantiating the accuracy and completeness of individual account transaction histories and account balances as of December 31, 2000 (see Section 4.3.3.1).

Response:

NORC designed the LSA sample to provide information regarding the accuracy of individual account transactions, as well as estimates of the total dollars in error. The completeness of the transaction histories is being tested in the Posting test, as described in the 2007 Plan.

Mr. Duncan - Page 5 of 79

The statistical sampling design employed attribute sampling, which is designed to answer yes or no questions such as was money collected properly recorded, and NOT to answer questions of "accuracy" such as how much money was collected (see Section 4.3.3.3).

Response:

The statistical sampling design did not employ attribute sampling. Mr. Duncan is incorrect in assuming that the dollar difference was not measured in the LSA reconciliation. It was. The sample was designed to measure the dollars in error and the dollar differences were measured by the LSA reconciliation. Therefore, there is no question of estimating the dollars in error from only an attribute. This statistical issue is described in more detail in Sections 2 and 4.

Mr. Duncan - Page 5 of 79

The 2007 Plan has a narrow and inappropriate definition regarding a deviation (or error), and hence a meaningless error rate. Both omitted transactions and transactions without supporting documentation are treated the same as properly supported, accurately recorded transactions (see Section 4.3.3.4).

Response:

Mr. Duncan's discussion of the definition of a deviation is not relevant because it assumes that no information is available for the dollar difference. Both the reconciliation status and the dollar difference are used in defining "error." As documented in the NORC LSA Report,

omitted transactions and transactions without supporting documentation are **NOT** treated the same as supported, recorded transactions. Mr. Duncan's theory as to how the errors were calculated is incorrect as described in Section 4 of my report.

Mr. Duncan - Page 5 of 79

The statistical sampling in the 2007 Plan does not constitute a sufficient basis for making any extrapolations beyond the population of recorded transactions in the Electronics Ledger Era subject to DOI's date restrictions (see Section 4.3.4.).

Response:

NORC has not extrapolated the results of the LSA beyond the LSA target population.   The sample which will be selected from the Paper Ledger Era transactions will provide a basis for estimation to that population. If the population of interest is expanded beyond that described in the 2007 Plan, the statistical sampling can also be expanded so that inference can be made to the larger target population. Section 3 of this report addresses these issues.

Mr. Duncan - Page 5 of 79

DOI's reliance on the Meta-Analysis Report to substantiate any assumptions or conclusions regarding data availability or data reliability is unfounded and inappropriate (see Section 4.3.5.).

Response:

Mr. Duncan has misunderstood the purpose and use of the Qualitative Meta-Analysis Report. Meta-analysis is included in the 2007 Plan as only one paragraph in Part 2. It is not central to NORC's sample design or estimation for the Paper Ledger Era, and to my knowledge, it is not central to the 2007 Plan. Additional responses to specific issues regarding the Qualitative Meta-Analysis Report are addressed in Dr. Scheuren's rebuttal report.

Mr. Duncan - Page 5 of 79

The statistical sampling in the 2007 Plan does not constitute a sufficient basis for making any extrapolations to accounts and/or transactions excluded by DOI from the historical accounting for accounts closed before October 25, 1994, direct pay transactions, deceased beneficiaries, and all transactions prior to June 24, 1938 (see Section 4.3.4.).

Response:

To my knowledge, the 2007 Plan does not attempt to make extrapolations to populations beyond the stated target. However, if the population of interest is expanded beyond that described in the 2007 Plan, the statistical sampling can also be expanded to the larger target population.

Mr. Duncan's Section 4. ANALYSIS

General Response to Section 4:

Mr. Duncan is mistaken in his assertions that the statistical inference has been extrapolated beyond the sampling population. In NORC's LSA Report, the LSA results are correctly applied to the population of 28+ million transactions from which the sample was drawn. The LSA Project covers a very large number of transactions in the population but it was not intended to cover all transactions in the Electronic Ledger Era. Mr. Duncan faults the 2007 Plan for omitting certain types of transactions from the target population. In fact, the 2007 Plan includes tests for portions of the Electronic Ledger Era transactions not covered by the LSA Project, as described in Section 3 of this report.

Mr. Duncan's report consistently contains his claim that

> *DOI's consultants in this matter have employed certain statistical sampling procedures that are critically flawed and would likely confuse and mislead the lay person. Furthermore, the results of the statistical sampling procedures are inappropriately applied, and hence will not support the DOI's stated objectives.*[14]

Mr. Duncan has strategically placed statements like this through his report. Mr. Duncan does not supply evidence that the sampling procedures are flawed and misleading. Mr. Duncan is incorrect in his statements regarding the statistical aspects of the LSA Project and the Plan. The statistical sampling and estimation procedures used are straightforward applications of standard statistical sampling to a target population. They are not critically flawed. The methodology is complex because the data we are dealing with are complex. While we do not expect a lay person to understand the complexities of the details of our analysis, NORC believes that it has explained the general concepts of the statistical analysis in a way that is understandable to a lay person. The overall standard for the analysis that NORC applied was to use established statistical techniques, i.e., ones that are well documented in the statistical literature and that will let the data speak.

My responses to specific statements in this section are provided below.

Mr. Duncan's Section 4.1.2. Land-Based Accounts

Mr. Duncan - Page 7 of 79.

DOI's statistical consultant, National Opinion Research Center ("NORC"), found that based on its sample results "…statistically valid conclusions can be drawn about all the land-based account transactions in the electronic ledger era."[17] Therefore, DOI claims to have completed the reconciliation work for Land-Based Accounts in the Electronic Ledger Era.

---

[14] Pages 13, 14, 16 and 31 of 79.

Response:

The quote is taken from the 2007 Plan but the conclusion Mr. Duncan draws in the sentence following is not correct when the entire Plan is taken in context. There is sufficient information in the 2007 Plan and in NORC's recommendations to refute this conclusion. I believe the quote from the 2007 Plan refers to NORC's recommendation that for the testing of the LSA target population, the sample size used in the LSA test was adequate and additional sampling from this portion of the population was not required. However, both NORC's recommendations and the 2007 Plan include additional work to be done in order to complete the testing for transactions in the Electronic Ledger Era. First, Dr. Scheuren's Expert Report, states that

> "follow-up work is required in order to make assurance statements about the error rates on transactions posted during the Electronic Ledger Era. There are two areas not yet covered:
> 1. *Posting tests* for the receipts in the Electronic Ledger Era - also referred to as completeness testing
> 2. *Clean-up tests* of transactions discovered later to be in the population of interest, but not included in the original sampling frame."[15]

The 2007 Plan includes both of these tests, as described in Part 2, Section V, "What Work Remains." On page 17, Part 2, the 2007 Plan states that the Land-to-Dollars posting test "addresses funds that should have been collected. A pilot land-to-dollars posting test has been completed for one region. Additional tests will be completed for each BIA region as part of the historical accounting."

On the following page in the 2007 Plan, a test of restored transactions is described as follows

> "The DCV and Interest Recalculation projects are designed, in part, to identify transactions and accounts that were missing from the electronic record. Once identified, these records are restored to the electronic record from historical system reports and financial documents. Because these transactions and accounts had not been identified in 2004, they could not have been selected in the sample drawn for the LSA reconciliation. **Interior plans to reconcile a sample of these restored transactions** to determine if they have an error rate significantly different from that found in the LSA sample. This work can only be performed as a follow-up test after all other work has been completed and the full population of additional transactions and accounts is known." (emphasis added)

Therefore, I believe the 2007 Plan clearly shows that DOI has not claimed to have completed tests for the Electronic Ledger Era.

Mr. Duncan – Footnote 16, Page 7 of 79.

DOI claims the total population of Electronic Ledger Era transactions under $100,000 is 28.8 million (23.2 million credit transactions and 5.6 million debit transactions).

Response:

There is no reference provided in this footnote, but the reference to 28.8 million that I am aware of is to the number of transactions under $100,000 that were in the target population for the **LSA Project**. As discussed in Section 3 of this Report, neither NORC in its LSA Report nor DOI have claimed that the LSA Project covers the entire target population

---

[15] Expert Report by Fritz Scheuren, August 2007, pages 7-8

Mr. Duncan's Section 4.1.2.1 Statistical Sampling of Land-Based Accounts

Mr. Duncan - Page 8 of 79.

It also appears that, despite minimal discussion in the 2007 Plan (a recurring feature of the plan), the "Meta-Analysis" performed by NORC plays a central role in purportedly strengthening and corroborating DOI's conclusions, including the results for the statistical sampling allegedly completed (i.e., Electronic Ledger Era), as well as currently contemplated (i.e., Paper Ledger Era). Given the significance placed on NORC's Meta-Analysis, the results and inferences are further discussed infra (see for example, Section 4.3.5.) as associated with specific areas of the 2007 Plan. EconLit's detailed analysis of NORC's Meta-Analysis is set forth in Appendix C.

Response:

The 2007 Plan included a paragraph on the Qualitative Meta-Analysis project in the Section titled "Lessons Learned to Date." This accurately represents the role played by this project. As stated in the 2007 Plan[16], "This work **suggests** that accounting efforts involving the paper ledger era will yield similar results as those found to date." (emphasis added) The important word is "suggests". No conclusions are drawn, but rather a test is being planned in order to draw conclusions.

Mr. Duncan's Section 4.2. Description of Statistical Sampling in the 2007 Plan

Mr. Duncan - Footnote 32, Page 10 of 79

These sample sizes purportedly represent the "in-scope" transactions. The number of "in-scope" transactions is less than the number of "original" transactions because somehow the "…reconciliation effort identified a small number of transactions as out-of-scope, i.e., they should not have been in the sampling population" (NORC LSA Report, p. 10). Rather than revisit the sample selection process to address the purportedly "out-of-scope" transactions, NORC and/or the accounting firms simply ignored them.

Response:

The out-of-scope transactions were not "ignored" but rather these transactions were fully documented and standard statistical techniques were used to account for them. The reasons for the out-of-scope transactions were discussed in Section 3 of NORC's LSA Report (page 8). Appendix B in NORC's LSA Report provided specific information on the transactions that were found to be out-of-scope and shows explicitly how the in-scope target population was estimated based on these transactions. Table 1 in the body of NORC's LSA Report shows the resulting estimated in-scope population. Section 3 in this report discusses the issues of coverage and out-of-scope transactions.

---

[16] Part 2, page 8.

Interestingly, the 2007 Plan does not contain any such assertions regarding extrapolating the results of the National Sample. Although no explicit language was provided in the 2007 Plan, EconLit understands that extrapolation of the sample results is clearly the intent. Clearly stated or not, DOI apparently intends to utilize the "results" of the National Sample, in conjunction with the previously discussed Meta-Analysis, in order to make the "accuracy and completeness" statements planned for the HSA's.

Response:

The point of this comment is unclear. Mr. Duncan appears to understand things that are not stated and of which I am unaware. The purpose of a statistical sample is usually to make estimates about the target population – an extrapolation. These statements cannot be made until the results of the sample are available, so results from planned tests cannot be provided prior to having the data. For the LSA Project, NORC provided estimates and upper bounds with 99% assurance statements for the target population covered by the LSA Project. These results were based on the data from the LSA Project only. The meta-analysis project was not a factor in the LSA results nor was it cited as such. In the future, NORC will continue to "let the data speak." This is the statistician's role. I expect that in the future NORC will be providing statements regarding accuracy and/or completeness and NORC will continue to make such statements based on the data resulting from statistically sound samples.

Mr. Duncan's Sections 4.2.1, 4.2.2, 4.2.3

General Response:

I would like to address two comments that Mr. Duncan makes repeatedly in the following three sections.

At the end of each section, Mr. Duncan places a statement such as the following (page 13 of 79)

Based on the information EconLit received, the Eastern Region statistical sampling procedures are critically flawed and would likely confuse and mislead the lay person. Furthermore, as discussed infra (see for example, Section 4.3.3.1), the results are inappropriately applied and will not support the DOI's stated objectives.

In the following three sections, Mr. Duncan does not state what he considers flaws in the statistical work. Rather Mr. Duncan attempts to cast doubt on the statistical abilities and professional standards of the NORC team by preceding statements from our reports with words such as "admits", "claims", and "purportedly". As one of the lead statisticians on NORC's team that provides statistical support services to OHTA, I disagree that the methods we have used are "critically flawed." The statistical sampling and estimation procedures used are straightforward applications of standard statistical sampling to a target population.

Mr. Duncan also repeatedly states that information was not provided to document what was actually done to reconcile transactions. For example on page 14 out of 78 he states that: *'it is not clear as to what was actually done to "reconcile" transactions.'* It is my understanding that the Accounting Standards Manual (ASM), which has been provided to the Court, is the manual describing how the reconciliation was done. The QA process was in place to ensure that the reconciliation was done to the standards laid out in the ASM.

I now turn to additional statements in these three sections.

Mr. Duncan's Section 4.2.1. The Eastern Region Sample

Mr. Duncan – Page 11 of 79

In other words, supporting documentation had already been identified for certain transactions and therefore transactions were "pre-selected" for inclusion in the sample. The number of "preselected" transactions in the original sample was substantial, consisting of 63 out of the 170 transactions with a value greater than $100, and 68 of the 95 transactions with a value less than $100.

Response:

While this was not the ideal sampling process it is not an unusual sampling situation. Sometimes the usefulness of sampling is not realized until a project has been started. NORC used appropriate statistical procedures in the estimation for the Eastern Region. I will describe in some detail here the issues and the methodology used, since it is not covered elsewhere in my report.

Because there are relatively few transactions in the Eastern Region IIM accounts, the original plan was to reconcile all transactions in this Region and the process of locating supporting documents was begun. Subsequently, it was decided that transactions would be sampled, but by that time many supporting documents had been located. Transactions where the documents were already collected were reconciled, but in computing estimates these "pre-selected" transactions would not represent any of the population other than themselves; in other words, it is not assumed that the "pre-selected" transactions are representative of the population. Therefore the "pre-selected" transactions would have a sampling weight of one. A random sample of the remaining transactions was selected.

All 170 transactions with a value greater than $100 were reconciled and so the fact that some were pre-selected and some were not makes no difference. When reconciled, no errors were found in either the 63 "pre-selected" or the remaining 107 transactions which did not have source documents on hand at the time of selection.

The "pre-selected" do make a difference in the strata where sampling is used. To keep the numbers in this example simple, suppose there are 400 transactions in a population and documents were already collected for 100. These transactions would be reconciled but they would not be representative of the remaining 300. A random sample would also be selected from the remaining 300. If a random sample of 30 is selected from the remaining 300, then the results from these 30 would be used to estimate the 300 and would each have a sampling

26

weight of 10. Suppose, in keeping with Mr. Duncan's worst case scenario, that the 100 "pre-selected" were in fact selected because they were the only accurate transactions out of the 400, i.e., the remaining 300 all have errors. In this overly simple example, the sample of 30 would all be found to have errors and the weighted estimate of the error rate would be calculated as 100*0 + 30*10*1 or it would be estimated that 300 of the 400 have errors. In the Eastern Region estimates, if errors had been found in the random sample, the errors would have been weighted similarly, to represent the errors in the portion of the population that was not "pre-selected." If on the other hand, both the 100 pre-selected and the 30 randomly selected had no errors, then the estimate is 0.

NORC used statistically appropriate procedures to ensure that the results would not be biased. The pre-selected transactions would only represent themselves, with a sampling weight of one[17]. For both the pre-selected and the randomly sampled transactions, the sample results were that no errors were found. In the end, weighting did not matter because no errors were found. There is no flaw in the statistical procedures here.

Duncan – Page 12 of 79

NORC admitted that the transactions found that were not in the original sampling frame indicate:

*"…that there was a **coverage problem** initially in using the electronic data file as the only sampling frame. The first problem is that not all eligible Eastern Region transactions were included in the electronic data – in other words that these are **data gaps**. The second problem is that some transactions on the original file appear to have been miscoded as 'interest' and therefore were excluded from the sampling frame" (emphasis added).*

The importance of this discovery by D&T and NORC will be further explored infra (see for example, Section 4.3.1.).

Response:

Sampling frames are rarely perfect. Omitted transactions or data gaps will be addressed by supplemental sampling later when these transactions are identified from the Data Completion Validation work. Section 3 of this report discusses these issues.

Duncan – Page 13 of 79

Based on the sample results, NORC makes the bold assertion that at a "…final sample size of 289, it is possible to make a 98+ percent assurance statement that the error rate is less than 1 percent, since no errors were found."[43] This assertion assumes the error rate used is appropriate for conducting an historical accounting. However, as discussed infra (see for example, Section 4.3.3.4), the error rates resulting from NORC's statistical sampling procedures are meaningless for conducting an historical accounting.

---

[17] NORC's Report "Eastern Region Sample Design and Selection", page 6.

Response:

The assurance statement made in NORC's report is a basic probability calculation and it is made only for transactions which can be reconciled with supporting documents, as stated in NORC's report immediately following this quote. Mr. Duncan has consistently misstated how error rates were actually calculated. The error rates are meaningful as I explain in Section 4 of my report.

Mr. Duncan's Section 4.2.2. The Alaska Region Sample

Mr. Duncan – Page 14 of 79

NORC uncovered a significant problem during the Alaska work. After "…the reconciliation was begun, it was determined that in the case of installment (or deferred) land sales, it was not feasible to simply reconcile in isolation one transaction out of the set of installment transactions."[46] NORC stated that this "…creates several problems for the data analysis and data summary."[47]

Response:

NORC did not "uncover" anything. As work is done, unanticipated situations arise and procedures need to be added or revised. The Accounting Standards Manual (ASM) is a "living" document. Most additions to the ASM have no impact on the sample estimation. However, when the accountants first encountered the installment land sales, in the Alaska pilot, it was determined that in order to measure the dollar difference, the sale had to be treated as one unit. Therefore if the sample selected one transaction in an installment land sale, the accountants would reconcile all transactions as a whole and determine the dollar difference for the entire sale. This increased the number of transactions to be reconciled and added unexpected complexity to the data collection. It also added complexity to the statistical estimation procedures. But this complexity was not a problem and standard statistical techniques were used, as described in Appendix C of NORC's LSA Report. Data complexity is not a "flaw."

Mr. Duncan – Page 14 of 79

Subsequently, NORC admits that "…we have inadvertently over-sampled the dollars in installment land-sales and particularly in sales with many payments. We cannot adjust for this by post-stratifying on the population because we cannot easily identify installment land sales by the transactional information alone."[48] Finally, NORC states that it "…also makes it difficult to use the measurement of the transaction difference rate, though one approach is to average the values over all the transactions."[49] The installment land sale problem also resulted in transactions that NORC deemed "out-of-scope" because there were land sales that did not conclude prior to December 31, 2000 (due to DOI's time period restrictions). Therefore, NORC removed a total of 22 credit transactions from the sample to be reconciled.

Response:

This is a further discussion of the reconciliation need to cluster transactions associated with installment land sales. As discussed in the previous response, it added additional complexity to the data base and to the estimation calculations. It also resulted in a lower than expected effective sample size for credit transactions. The quote from NORC's report points out why one standard statistical estimation technique for improving precision could not be applied in this case. Again, complexity is not a flaw.

The out-of-scope transactions are described in a later comment in this section and more generally in Section 3 of this report. I want to emphasize that the definition of out-of-scope was determined and recorded as part of the reconciliation process and was not solely NORC's prerogative.

Mr. Duncan – Page 14 of 79

After removing the 22 credit transactions for the installment sale issue, NORC claims to have "reconciled" 418 transactions out of 423 low dollar value transactions (i.e., 239 of 242 debit transactions and 179 of 181 credit transactions). NORC also claims that no differences were found among the reconciled debit transactions. However, for the credit transactions, NORC states that "…differences were found in both the high dollar and sampled portions of the credit populations. However, because of the large number of installment land sales in Alaska and the small sample size, in this report we provide only descriptive statistics and do not attempt to make inferential statements from the sample of credit transactions" (emphasis added).[50]

Response:

First I want to clarify that no one on the NORC team reconciled transactions. As described in Section 4, NORC analyzed the data returned from the reconciliation process.

The Alaska Report was a report on a pilot study. One of the things learned from the pilot study was that installment land sales must be reconciled as a group. Land sales are much more common in Alaska than in the other Regions, and therefore this determination regarding land sales had the effect of noticeably reducing the effective sample size for the credit transactions in the Alaska Region. The other effect was to increase the cost of reconciliation. In one example, the 19 selected transactions could not be reconciled without reconciling an additional 287 transactions. Had the purpose of the LSA sample for Alaska been to provide strong assurance statements regarding the error rates for Alaska, I would have recommended that the sample size be increased, i.e., select and reconcile additional transactions. However, that was not the purpose of the Alaska sample for the LSA Project, and so no additional transactions were selected. Again this is not a critical flaw for the LSA Project.

Mr. Duncan – Page 14 of 79

Despite the actual conclusions in the NORC Alaska Analysis Report, the NORC LSA Report incorporates the original sample size of 442 (i.e., 445 less 3 out-of-scope) into the sample of

4,500 low dollar value transactions. Furthermore, the purportedly reconciled transactions for the Alaska Region are reported as 437 transactions out of 442 transactions in the NORC LSA Report (as compared to the 418 transactions out of 423 transactions in the underlying NORC Alaska Analysis Report). No explanation was provided regarding the changes to the Alaska Region sample, or for the inconsistencies in results between the NORC Alaska Analysis Report and what was represented in the 2007 Plan from the NORC LSA Report.

Response:

The Alaska Region sample represents one stratum in the stratified sample over all 12 BIA Regions so of course it was included in the final estimate. Stratified sampling was used to ensure coverage over all Regions but not to provide precise estimates for each Region. It would have been a critical flaw if we had not included these sampled units in the final estimate.

I will agree that the NORC LSA Report should have clarified this difference between the two reports. As discussed above, the determination that all transactions resulting from an installment land sale must be reconciled together, as one "event", resulted in several additional complexities to the process. When an installment land sale included both transactions that occurred prior to December 31, 2000 and transactions that occurred subsequent to December 31, 2000, the question arose as to whether such a cluster of transactions should be included in the population of transactions through December 31, 2000, or not. The 22 transactions considered out-of-scope at the time the Alaska Report was written (2004) included 19 transactions in one installment land sale which was not completed until after December 31, 2000. The Alaska reconciliation was a pilot and the Alaska Report was written much in advance of the final LSA Report. When the Alaska Report was written, it had been provisionally decided to define these 19 transactions as out-of-scope. On further consideration, it was decided that a better approach was to include such transactions in the target population, and reconcile such sales. Valid estimates could be calculated by appropriately attaching any differences to the selected transactions, as described in Appendix C of NORC's LSA Report. Between the Alaska Report and the LSA Report, the 19 transactions were reconciled.

Mr. Duncan's Section 4.2.3. The 10-Region Group Sample

Mr. Duncan – Page 15 of 79

The accounts and transactions were divided into 4 "replicates" to be purportedly reconciled by 4 different accounting firms. The final transaction count for low dollar value sampling of the 10- Region Group was 20,357 transactions (with each of the 4 "replicates" containing between 4,932 and 5,203 transactions). Despite what appears to be a substantial and complicated effort to get to the sample of 20,357 for the 10-Region Group, the Office of Historical Trust Accounting ("OHTA") subsequently made the unilateral decision to reconcile "…only Replicate 1…"[53] Therefore, apparently the "sample weights" were somehow "recalculated" to "…reflect the fact that only approximately one-fourth of the original National Sample was to be reconciled."[54] The final transaction count for the low dollar value sample was 5,138 transactions.

Response:

Mr. Duncan misunderstands the use of replication and appears to be unfamiliar with how sample weights are calculated. The use of replicates in sample design is a standard statistical technique. By randomly dividing the sample into equal parts, i.e., replicates, and reconciling one replicate at a time, statistical estimates can be made prior to the completion of the entire sample. When the elements in the first replicate are completed, valid estimates can be made using only the first replicate, as it is a valid random sample. As additional replicates are completed, the estimates become more precise as the sample size increases.

For the LSA Project, there was a concern that the time and resources needed for performing the reconciliation might be under-estimated. If the entire sample was started (rather than just the first replicate) but could not be completed with the available resources, usable inference from the sample might not be possible. NORC recommended that the LSA Project begin by first reconciling one replicate. As the work progressed, it became clear that the reconciliation was more resource intensive than expected, and the decision was made by DOI to complete only the first replicate.

The calculation of replicate weights is a straightforward calculation. The sample weight represents the number of population units represented by the sample unit and is the inverse of the probability of selecting the unit. Replicates are formed using probability sampling and the calculation of the replicate sample weight uses a straightforward probability calculation. If the sample is divided into four replicates, the sample weight for each unit in the first replicate is approximately four times the original sample design weight. For example, suppose the original sample selects 40 units out of 400 population units, then the probability of selection for each sample unit is 40/400 and the original design weight is 400/40 or 10. If four replicates are formed, when the estimation is calculated based on the first replicate, the sample will contain 10 units randomly selected out of the 400 population units and the sampling weight will be 400/10 or 40. The weights are calculated using the known probabilities of selection.

Mr. Duncan – Footnote 56, Page 13

NORC did provide a "comparison of reconciliation results between 2004 and 2005" section in the NORC LSA Report (see pp. 11-12). It is notable that in "…2004, among the 1,429 credit transactions reconciled under $100,000, there were 11 found to have differences (a rate of 0.8%). Out of the 330 credit transactions reconciled in 2005, 19 differences were found (a 5.8% rate). This is more than a seven-fold increase" (see NORC LSA Report, p. 12). NORC further admits that the difference between the 2004 reconciled transactions and the 2005 "subsample" is "…highly statistically significant…" (see NORC LSA Report, p. 12).

Response:

Mr. Duncan does not go further to note that NORC's final estimates reflect this difference by correctly giving greater weight to the subsample results. Therefore this does not describe a "flaw." The correct, probability-based weights were used for the subsample, which now also represent the original sample units that were not reconciled in 2005. That is, the sampling

weights for the subsampled units reflect the additional subsampling process and the difference between the two groups is reflected correctly in the final estimates.

Mr. Duncan's Section 4.3

General Response to Mr. Duncan's Section 4.3:

In Section 4.3 of Mr. Duncan's report, *4.3 Analysis of Statistical Sampling in the 2007 Plan* he lists out his specific, statistically related problems with the 2007 Plan. His five key problems with the 2007 are:

1. Data problems;
2. Sample selection problems;
3. Sample design problems;
4. Extrapolation problems; and
5. Meta -analysis problems.

Much of the criticism in the following section depends on Mr. Duncan's mistaken claim that the statistical results of the LSA Project have been applied to the entire population of transactions in the Electronic Ledger Era. Section 3 in my report discusses why this claim is not true and outlines how additional tests described in the 2007 Plan address other portions of the population, the "gaps". Much of his discussion in this section depends on this incorrect assumption on his part.

Since many of the points raised are discussed in Sections 2-4 of my report, I have kept my additional responses to a minimum.

Mr. Duncan's Section 4 .3.1. Data Problems

Mr. Duncan – Page 17 of 79:

Plaintiffs have documented fundamental failings of both the historical transaction records as well as potential supporting documents. For example, Plaintiffs assert that the relevant data is unreliable because it has been subject to fraud, adulteration and misappropriation,[58] and that the relevant data is incomplete because many necessary documents have been accidentally or deliberately destroyed.[59]

Response:

As I have previous said, NORC's approach is to let data speak for itself. The LSA target population does not address the data gaps (see Section 3) but it does address a target population of 28 million transactions which were recorded in the IRMS and TFAS data base. The LSA results indicate that supporting documents were located to allow the reconciliation process to determine the dollar differences for over 99% of the transactions in the sample. Errors were found in the sampled transactions, but NORC's statistical estimates based on the

data from the reconciliation process indicate that the state of these transaction records is much better than what Plaintiffs have asserted.

Mr. Duncan – Page 18 of 79:

*The starting point for the LSA project was a recorded transaction, and "…any failure to collect, deposit, and record collection transactions would likely not have been discovered in LSA project testing."[69] In fact, as previously discussed, NORC actually experienced and documented exactly this issue (for the Eastern Region, where enough detailed information was provided). There were 11 transactions that were found in the Eastern Region sample that were not in the electronic database used for sampling (i.e., the "data gaps"), of which 2 could not be reconciled (even by the unknown "alternative procedures" method). NORC clearly stated that "[t]wo of these transactions have not been reconciled and these two missing transactions without supporting documents are particularly troubling, as it could be argued that such transactions pose the **greatest risk of errors**" (emphasis added).[70]*

Response:

This paragraph mixes two ideas. First, regarding the issue of transactions not recorded in IRMS or TFAS, this example would appear to illustrate that such transactional information can be obtained from other government sources, in this case, possibly supporting documents. As stated in the 2007 Plan, there is a follow-up test where transactions such as these will be tested.

The issue of whether it is less likely that such transactions can be reconciled is still an open question. The test has not been performed yet, but such a test is contained in the 2007 Plan. In the LSA Project, there were transactions where not all supporting documents were located. But the LSA Project results indicate that for the LSA target population, such transactions were few and for estimation, NORC counted these transactions as "errors." Until we complete the follow-up samples for the LSA Project we will not know whether it will be more difficult to locate the supporting documents for the transactions that were initially not included in the data base. But as described in Section 4, NORC is well aware of the statistical issues involved with incomplete data.

Mr. Duncan – Page 19 of 79:

*The Eastern Region Sample also provided an important piece of evidence regarding missing documents that casts a shadow on the entire "reconciliation" process. As set forth in Table 4-1: Eastern Region Sample Results, 15 percent of the sample was "reconciled" under "accounting code 2". EconLit understands that "accounting code 2" indicates that no contemporaneous documentation could be found (i.e., unable to obtain directly relevant documents, unable to recalculate using known DOI business practice, or unable to obtain third-party confirmations)[71], and therefore "alternative procedures" were employed to purportedly reconcile these transactions. This is further support that there are significant numbers of transactions without supporting documents. However, it is important to note that DOI still considered these transactions **<u>reconciled with no errors</u>**. A further discussion of the reconciliation framework, including the "levels" of supporting documentation, with reference to Plaintiffs' accounting expert, is discussed in Section 4.3.3. and Appendix D.*

Response:

This is not my understanding of the meaning of accounting code 2 and I would refer Mr. Duncan to the ASM. The transactions with accounting code 2 are considered as reconciled in NORC's estimation but these transactions can have dollar differences, which would be counted as errors in the estimation. I have also provided a response to errors in the material in Mr. Duncan's Appendix D.

Mr. Duncan – Page 19 of 79:

Additionally, NORC had previously reported on the vast quantity of missing data in the Electronic Ledger Era for all 12 BIA Regions with respect to the conversion to the Integrated Records Management System ("IRMS").[72] EconLit's analysis and summary of NORC's findings on the number of months of missing data found that between 1985 and 1999 an average 4.5 percent of the data from each agency was missing from IRMS.[73] In fact, some agencies are missing as much as 12 years of data between 1985 and 2000. This is further support that months of data for thousands of accounts is missing from IRMS, and therefore from the population sampled by the LSA project.[74]

Response:

As described in Section 3, the LSA Project never claimed to cover these transactions, but there is a plan to also test these transactions.

Mr. Duncan – Page 19 of 79:

Missing and/or destroyed data in the Paper Ledger Era is also significant. The data availability in the Paper Ledger Era is such that DOI and NORC have not been able to adequately identify the population of accounts and transactions from which to sample.

Response:

Mr. Duncan confuses "no information available *electronically*" with "no information." By definition, the Paper Ledger Era information is not currently available as an electronic data base. Therefore the sample cannot be designed in the same way as the Electronic Ledger Era, but statistically valid samples do not have to rely on electronic data bases. Information is currently available to develop the sampling frame to cover the population of accounts with Paper Ledger Era transactions. The transactional data are not currently available electronically, but can be obtained from paper documents.

Mr. Duncan's Section 4.3.2. Sample Selection Problems

General Response to Mr. Duncan's Section 4.3.2:

This section depends entirely on Mr. Duncan's assertion that the estimates from the LSA Project were applied to the entire population of transactions in the Electronic Ledger Era. As discussed at length in Section 3, this is incorrect. The estimates were not extrapolated

beyond the target population for the LSA Project and therefore Mr. Duncan's arguments in this section do not apply. The statistical inference was applied correctly.

Mr. Duncan – 1st paragraph, page 22 of 79:

Problems associated with differences between a sampled population and the target population cannot be addressed by increasing the sample size or any other statistical procedure that doesn't address drawing a new sample from the target population. In this instance, drawing a sample from the target population is not possible.

Response:

I have already addressed the issues that Mr. Duncan describes as problems in Section 3, but let me reiterate, the LSA sampling frame appropriately covers the target population that was defined for the LSA Project. Therefore, our sample, which is a valid sample from our sampling frame, is a valid sample for the target population. Additionally, all estimates and confidence bounds that NORC has calculated based on the samples, are appropriate and validly represent the target population. Mr. Duncan is apparently defining a different target population.

Mr. Duncan's Section 4.3.3.1 Irrelevant Sample Design

General Response to Mr. Duncan's Section 4.3.3.1:

Mr. Duncan's arguments hinge on a false assumption that the error rates DOI calculated "netted" the errors. In Section 4 of this report I explain how errors were calculated and accumulated in the LSA Project and there was no netting of errors. Section 4 provides my response to this section and therefore I have kept my specific responses here to a minimum.

Mr. Duncan – Page 23 of 79

By way of example, if two transactions in different accounts are misstated, one an over-payment of $100 and the other an under-payment of $100, the mean under-payment[78] according to DOI will be zero even though both of the transactions are misstated and both of the account balances are wrong. Because of this problem, the 2007 Plan will be unable to provide any level of assurance to individual account holders. Based on the sample design in the 2007 Plan, the only conclusion that can be drawn from a low mean under-payment is that for mistakes in a given individual account, an offsetting transaction accrued to the benefit of somebody else, somewhere else, at some other time.

Response:

Mr. Duncan makes several claims, on page 23 of 79, such as this. The claim is that the estimates of error allow overpayments and underpayments to cancel out. This is incorrect. As described in Section 4 of this report, differences were not netted across transactions.

*The statistical sampling procedures were designed for **estimating litigation exposure** and NOT for substantiating the accuracy and completeness of individual account transaction histories and account balances as of December 31, 2000 (DOI's stated objective).*

Response:

As discussed in Section 2, NORC designed the sample with the objective of providing estimates and inference about the accuracy of transactions.

Mr. Duncan's Section 4.3.3.2 Inappropriate Sampling Unit

Mr. Duncan – Page 23 of 79:

*Notwithstanding that the sample design is fundamentally irrelevant, the 2007 Plan retains a critical flaw of the 2003 Plan by sampling **transactions rather than accounts**. In other words, the 2007 Plan "reconciles" transactions and makes statements about the accuracy and completeness of accounts and account balances. The 2007 Plan assertion regarding accurate **account balances** based on the accuracy rate of **transactions** is unfounded and incorrect. Assuming that the DOI could sample from the target population (which EconLit disputes elsewhere) and assuming that accurate opening account balances were established, statistical inferences based on transaction sampling that could be made about transactions are not equivalent to statistical inferences that could be made about account balances.*

Response:

As Mr. Duncan points out, the 2007 plan defines the historical accounting as follows:

> For purposes of this 2007 Plan, the historical accounting is the provision to each IIM account holder of his/her account transaction history in an HSA—a listing of all transactions in an IIM account—and a separate statement regarding the **accuracy and completeness of account transactions**.[18] (emphasis added)

The 2007 plan clearly states that an accuracy and completeness statement will be provided with respect to transactions, not with respect to account balances. Mr. Duncan is correct that the opening section of the plan states "In addition, Interior plans to provide each IIM account holder with Interior's conclusions about the accuracy of the account transaction history and the account balance as of December 31, 2000." However, NORC was never asked to design a sample which would provide the individual with a statistical statement about his or her account balance. This would appear to be an example where Mr. Duncan states what he believes should be the objective of the plan, and how the plan does not address his objectives.

---

[18] 2007 Plan, Part 1, p. 9. "HSA" refers to the "historical statement of account" that DOI proposes to provide IIM account holders.

The LSA sample will enable DOI to make statements regarding the accuracy of account transactions processing.

Mr. Duncan – Page 24 of 79:

By way of example, statistical inference based on the assumed 99 percent accuracy rate of transactions suffers from the following limitation:

*The average land-based account open on December 31, 2000 had approximately 137 transactions.[79] Assuming the 2007 Plan assertions regarding a 99 percent transaction accuracy rate are correct, each individual transaction has a 99 percent chance of being correct and a 1 percent chance of being wrong. When the probabilities are applied to the average number of transactions per account (137) there is only a 25 percent[80] probability that all of the transactions in the account are correct and a **75 percent probability that at least one erroneous transaction exists**.[81] Therefore, ignoring the other fundamental flaws in the 2007 Plan, **there is a 75 percent probability that the account balance is misstated**. Any reasonable assurance statements regarding IIM account balances based on the "accuracy rate" of transactions are unfounded and incorrect.*

Response:

NORC has not been asked to provide statistical assurance statements regarding IIM account balances from the LSA sample results. Again, the objective of the LSA sample was to assess the accuracy of account transactions. If one further wished to make statements about accounts or account balances, the estimation and calculation issues are complex, and rely on assumptions. For example, the calculation that Mr. Duncan makes depends on an assumption that the errors are independent. However, one could argue that certain types of errors, such as errors in the allocation by the ownership, might not be independent.

Mr. Duncan's Section 4.3.3.3 Inappropriate Type of Sample Plan

Mr. Duncan – Page 25 of 79:

The sampling proposed in the 2003 Plan was explicitly based on attribute sampling of certain Land Based IIM accounts. The Duncan Report and the Duncan Rebuttal Report specifically pointed out the shortcomings of designing an attribute sample for the stated purpose of the 2003 Plan. Rather than directly address or supplement the fundamental shortcoming in the sample design within the 2003 Plan, the 2007 Plan simply removes all references to attribute sampling. It is, however, clear that the 2007 Plan is still based on an attribute sample design.

Response:

There is no reference to "attribute sampling" in the 2007 Plan because it is not used in the Plan. The entire premise of this section is that the sample was designed to only estimate the "attribute." As discussed in Section 2 of this report, this premise is incorrect. The LSA Project measured and reported the dollars in error for each reconciled transaction. The LSA

Project was not an attribute sample design and, to the best of my knowledge, the reconciliation in the Posting Test and in the Paper Ledger Era will also measure dollar differences.

Mr. Duncan –Page 25 of 79:

The NORC appendix to the 2003 Plan which discussed how an estimated maximum dollar error might be computed includes a reference to the Audit Sampling book co-authored by Dan. M. Guy ("Guy Text").[88] The Guy Text states, "Three techniques of limited accuracy to translate attribute information into dollar estimates are discussed below. Before studying these techniques, remember that the primary objective of attribute sampling is not dollar estimation."[89] The primary objective of attribute sampling is to estimate the rate of occurrence of a specific quality or attribute in a population,[90] not to extrapolate dollar implications to the population.

The citation Mr. Duncan selects from the textbook by Dan M. Guy is a reference to a section discussing the case when only the yes/no attribute is measured and recorded. Basically it says that it is very difficult to estimate dollars in error if the sample does not measure dollars in error. The LSA Project measured and reported the dollars in error for each reconciled transaction. Therefore this citation is not applicable.

Mr. Duncan - Page 26 of 79:

Therefore, the 2007 Plan relies on an inappropriate statistical sampling methodology that is not designed to accomplish the 2007 Plan's stated objectives.[92]

Response:

To summarize, the LSA Project measured the dollar difference and it is my understanding that the reconciliation described in the 2007 Plan will also measure dollar differences, not simply the attribute. Therefore the statistical sampling methodology is designed to accomplish the stated objectives.

Mr. Duncan's Section 4.3.3.4 Overly Narrow Definition of "Errors"

General Response to Mr. Duncan's Section 4.3.3.4:

Mr. Duncan's discussion in this section relies on the incorrect assumption that the magnitude of the dollar difference is not measured in the reconciliation. This assumption is not correct and Sections 2 and 4 of my report address these issues. For the LSA Project, the observed dollar difference is recorded in the data base and it is NORC's understanding that the dollar difference will continue to be measured as part of the reconciliation. The magnitude of the dollar difference is measured and therefore it is discovered. Mr. Duncan's discussion here of attribute sampling is not relevant.

His discussion of how errors are counted also appears to be assuming that transactions not in the target population are included in the estimation. The LSA Project results were applied only to the specific target population for that project and not to "omitted" transactions. As

described both in the NORC LSA Report and in Section 4 of this report, NORC counted unreconciled transactions as errors.

Mr. Duncan – Page 28 of 79:

An additional complication in the 2007 Plan's treatment of errors is the offsetting of negative "errors" with positive "errors" to arrive at the "net error rate". The result of calculating a "net error rate" is a material misstatement of the true underlying errors experienced in individual accounts. Therefore, the statements made in the 2007 Plan regarding "small net error rates" are misleading and ultimately irrelevant to individual Indian account holders.

Response:

As explained in Section 4 of this report, the LSA Project results did not use any such "netting" of errors as described here. I see no evidence that this would occur in the future.

Mr. Duncan's Section 4.3.4. Extrapolation Problems

Mr. Duncan – 3rd paragraph, page 29 of 79:

However, as previously stated, although the 2007 Plan contains no explicit language regarding extrapolating the results of the National Sample, EconLit understands that extrapolation of the sample results is clearly the intent. Accordingly, as was the case with the 2003 Plan, the statistical sampling in the 2007 Plan does not constitute a sufficient basis for the proposed extrapolation (implied or otherwise).

Response:

Mr. Duncan cannot provide a basis for his "understanding" of the "intent" so it is difficult to know what proposed extrapolation he refers to. I believe that Mr. Duncan is referring to extrapolating a sample beyond the population from which it was drawn. This was not done. Section 2 of my report addresses this issue.

Mr. Duncan – Page 29 of 79:

Additionally, EconLit understands that Plaintiffs have asserted that many accounts and/or transactions have been inappropriately excluded by DOI from the historical accounting mandate. Such exclusions include accounts that have been closed before October 25, 1994, direct pay transactions, deceased beneficiaries, and all transactions prior to June 24, 1938. The statistical sampling in the 2007 Plan does not constitute a sufficient basis for making any extrapolations to these accounts and/or transactions that have been excluded by DOI from the historical accounting.

Response:

The 2007 Plan addresses a very specific population of accounts and transactions. The sample would be designed for this population. However, if it is determined that the target population

should be expanded beyond that in the 2007 Plan, the sample design can also be expanded to cover a revised target population.

Mr. Duncan's Section 4.3.5. Meta-Analysis Problems

Mr. Duncan – Page 29 of 79:

The Meta-Analysis performed by NORC plays a central role in purportedly strengthening and corroborating DOI's conclusions, including the results for the statistical sampling allegedly completed (i.e., Electronic Ledger Era), as well as currently contemplated (i.e., Paper Ledger Era). EconLit's detailed analysis of NORC's Meta-Analysis is set forth in Appendix C.

Response:

I have previously explained that the Qualitative Meta Analysis results did not play a central role in DOI's conclusion. NORC's Dr. Scheuren is preparing a separate rebuttal to the accusations Mr. Duncan makes concerning NORC's Qualitative Meta-Analysis Report.

Mr. Duncan's Appendix D. Accounting Standards Manual: Reconciliation Accounting Code

General Response to Mr. Duncan's Appendix D:

In this appendix, Mr. Duncan has drawn his own conclusions as to how differences were counted. In most cases, he is incorrect.

First, there is no process rule that automatically assigns an "error" to a transaction. For each transaction, the reconciliation provides data on the reconciliation status, and if reconciled, the dollar difference found. These data are on the data base and can be used to determine what should or should not be counted in the estimation. Mr. Duncan's implications in this Appendix as to how errors were counted are not correct.

When the transaction is reconciled and there is a dollar difference found of $1 or more, the transaction is considered to be 'reconciled with an error'. Therefore his diagrams for accounting code one and accounting code two are incorrect.

My understanding from the ASM is that accounting code 3 indicates partial reconciliation. There is no process rule for how this should be counted. NORC's estimation used the rule that the designation of "error" would depend on whether or not there was a dollar difference reported from the work that could be done. Therefore, NORC counted a transaction with accounting code 3 and a dollar difference as 'reconciled with an error.' A transaction with accounting code 3 and no dollar difference was counted as 'reconciled and no error.'

Accounting code 4 is not considered reconciled and therefore there is no information regarding whether or not there is a dollar difference. As described in NORC's LSA Report, for estimation purposes NORC did count these transactions – but we counted them as errors. Therefore Mr. Duncan's assumption is also incorrect in this diagram.

This issue is discussed in more detail in Section 4 of this report.

Mr. Duncan's Appendix E. Eastern Region Transaction Reconciliation

General Response to Mr. Duncan's Appendix E:

In this appendix, Mr. Duncan displays graphically the Eastern Region results by reconciliation code. His last chart is a misrepresentation of the information supplied in the cited reference. This reference explicitly lists the accounting code and the dollar differences found for each reconciled transactions. Therefore, the information is provided which shows for the transactions in the sample, the reconciled transactions with supporting documentation that matched the dollar amount of the transactions. The NORC report provides an extrapolation of the sample results to the target population.

Appendix B. Compensation

September 17, 2007

<u>Statement of Susan Hinkins, PhD</u>

Re: <u>Cobell v. Kempthorne</u>

I submit the following statement regarding my compensation in connection with service as an expert in this matter: I am employed through NORC, which is under contract with the Department of the Interior. NORC receives in compensation for my work the hourly rate of $226.48. Neither NORC nor I am being separately compensated for this report or for any testimony I may give.

_____

Susan Hinkins, Ph.D.

Appendix C. List of Sources/References

Expert Report of Dwight D. Duncan, August 23, 2007

Expert Report of Dr. Fritz Scheuren, August 2007

Historical Accounting Project, Part 1, Plan for Completing the Historical Accounting of Individual Indian Money Accounts, DOI, May 31, 2007

Historical Accounting Project, Part 2, The Basis and Rationale for Changes to the January 6, 2003 Historical Accounting Plan for Individual Indian Money Accounts, DOI, May 31, 2007

Reconciliation of the High Dollar and National Sample Transactions from Land-Based IIM Accounts (All Regions), NORC, 2005

Litigation Support Series, Alaska Region Sample Design Report, NORC, 2004

Litigation Support Series, Analysis of the Alaska Sample, NORC, 2004

A Statistical Evaluation of Preliminary Eastern Region Sample Results, NORC, 2004

Eastern Region Sample Design and Selection, NORC,2003

Accountants Report on the Reconciliation of the Eastern Region Land-Based Non-Interest Individual Indian Money Transactions, Deloitte & Touche, 2003 NORC Sample Design Planning Report, NORC, 2003

Accounting Standards Manual

Thompson, S. (1992) *Sampling*, John Wiley & Sons

Guy, D.M., Carmichale, D., Whittington, O.R. (1998), *Audit Sampling*, John Wiley & Sons

Appendix D. Resume for Susan Hinkins

EDUCATION

Ph.D.   Statistics, Montana State University, Bozeman, 1979

M.S.   Mathematics, Montana State University, Bozeman, 1973

B.S.   Mathematics, University of Wisconsin, Madison, 1971

PROFESSIONAL EXPERIENCE

2001-present   Senior Statistician, National Opinion Research Center

Dr. Hinkins joined NORC in December 2001, where she joined the team working on the historical accounting of the Individual Indian Money (IIM) accounts, for the Department of the Interior and in September 2002 she became the project manager. The purpose of the historical accounting is to provide the individual Indian trust fund beneficiaries with information that will allow them to ascertain whether the Secretary has faithfully fulfilled the IIM trust. NORC's role, generally, is to assist in the planning and implementation in the steps for an accounting in a manner that allows the measurement of the confidence in the results. Samples are developed in order to test the accuracy and completeness of different pieces of the process. NORC helps with determining frames, developing sampling plans, overseeing the selection of random samples, and writing the reports to provide the needed administrative record of the project activities. NORC also designed a customer satisfaction survey for the Office of Indian Trust Transition and developed cognitive interviews for pre-testing the survey.

1998-2001   Manager, Quantitative Economics and Statistics Group, Ernst & Young,LLP.

Dr. Hinkins' responsibilities included general statistical consulting with an emphasis on estimation from complex data structures. Projects included the design and analysis of a survey of members of a business related organization; writing expert witness testimony to refute the results of a survey; designing samples to estimate the accuracy of claims; designing a matrix sample to estimate inventory proportions with a small sample and using replicate variance estimates; estimating treatment differences from data collected from an observational study; and designing a sample to investigate the accuracy of tax records.

Outside of her employment, she testified as an expert witness on a Title IX discrimination case in U.S. District Court and she was a member of a team advising the South African Finance Department on how to develop data for use in revenue and economic models.

1981-1998   Mathematical Statistician, Statistics of Income Division,
             Internal Revenue Service.

Responsibilities included sample design and estimation issues for complex samples, primarily for the annual sample of corporate tax returns. Projects included: responsibility for calculating advance estimates, before the sample was complete; devising resampling and replication methods for point estimation and variance estimation that would be more accessible for users; developing and maintaining a double sampling procedure to collect certain information for only a subset of the sample units - the missing information was estimated using a hot deck multiple imputation technique.

1986 - 1989   Mathematical Statistician, Biology Department, Montana State University.

Mathematical statistician on a project funded by the Environmental Protection Agency to do exploratory data analysis of a stratified sample of lakes in the U.S. The survey was to provide baseline measurements for the water quality in U.S. lakes.

1980 - 1981   Mathematical Statistician, Office of Radiation Programs, U.S. EPA

Coordinated a study to compare the reliability and validity of various methods for measuring the level of radon and radon daughters in homes, coordinating the work between the D.C. office, two EPA labs, and statistical consultants at two universities.

1976 - 1980   Research Assistant and Statistical Consultant, Montana State University

Statistical consultant for various research projects for the Fisheries Bioassay Laboratory and the U.S. Forest Service, as well as providing statistical assistance on graduate projects.

SELECTED PUBLICATIONS

Discussion of "Undoing Complex Survey Data Structures: Some Theory and Applications of Inverse Sampling" by J.N.K. Rao, A.J. Scott and E. Benhin, *Survey Methodology,* 2003.

"Application of Matrix Sampling Design on Inventory Estimation," with Robin Lee and John Matson, *Proceedings of the American Statistical Association*, 2001.

"Application of an Inverse Sampling Algorithm to a State-Level National Health Interview Survey," with Fritz Scheuren and Van Parsons, *Proceedings of the American Statistical Association*, 1999.

"Design Free Analysis," with Fritz Scheuren and Yan Liu, invited paper presented at the Statistical Society of Canada meeting, June, 1998.

"Inverse Sampling Design Algorithms," with H. Lock Oh and Fritz Scheuren, *Survey Methodology*, 1997.

"Replicate Variance Estimates - Reducing Bias by Using Overlapping Replicates," with H. Lock. Oh and Fritz Scheuren. *Proceedings of the Section on Survey Research Methods, American Statistical Association*. 1997.

"Replicate Variance Estimation in Stratified Sampling with Permanent Random Numbers," with Chris Moriarity and Fritz Scheuren, *Proceedings of the American Statistical Association*, 1996.

"Creation of Panel Data from Cross-Sectional Surveys," with Stephanie Hughes, *Proceedings of the American Statistical Association*, 1995.

"Inverse Sampling Design Algorithms," with H. Lock Oh and Fritz Scheuren, *Proceedings of the American Statistical Association*, 1994.

"Statistics of Income Division's Uses of Administrative Business Tax Records: An Overview," with Jeri Mulrow and Jonathon Shook, *Proceedings of the International Conference on Establishment Surveys*, 1993.

"Comparing Advance and Final Estimates: 1990 SOI Corporate Sample," with John Czajka, *Proceedings of the American Statistical Association*, 1993.

"Evaluating Sample Design Modifications: Balancing Multiple Objectives," with Fritz Scheuren, *Proceedings of the American Statistical Association*, 1989.

"Updating Tax Return Selection Probabilities in the Corporate Statistics of Income Program," with Homer Jones and Fritz Scheuren, *Proceedings of the Statistics Canada Symposium on Statistical Uses of Administrative Data*, 1987.

"Hot Deck Imputation Procedure Applied to a Double Sampling Design," with Fritz Scheuren, *Survey Methodology*, 1986.

Discussion of "Survey Nonresponse Adjustments," by Rod Little, *Proceedings of the American Statistical Association*, 1984.

"Imputation of Missing Items on Corporate Balance Sheets," *Proceedings of the American Statistical Association*, 1982.

"RFACTOR - A program to Create Rubin's Factorization When there are Incomplete Multivariate Data," *American Statistician*, 1980.

PROFESSIONAL ACTIVITIES

Fellow of the American Statistical Association

Member of the Institute of Mathematical Statistics

President of the Montana Chapter of the ASA 2005; Vice-President 2004; Secretary/Treasurer 1997-1999.

Treasurer of the Social Statistics Section 2005-2006.

Chair of the ASA Committee on Scientific Freedom and Human Rights 2005-2007

Refereed papers for JASA, American Statistician, Survey Methodology, and the Journal of Official Statistics

Pro bono expert witness for the plaintiffs in a Title IX discrimination case in U.S. District Court

Member of a team advising the South African Finance Department on how to develop data for use in revenue and economic models

Appendix E. Acronyms

ART	Account Reconciliation Tool
ASM	Accounting Standards Manual
BIA	Bureau of Indian Affairs
DCV	Data Completeness Validation
DOI	The United States Department of the Interior
IIM	Individual Indian Money
IRMS	Integrated Records Management System
LSA	Litigation Support Accounting
NORC	National Opinion Research Center at the University of Chicago
OHTA	Office of Historical Trust Accounting
TFAS	Trust Funds Accounting System